Lineaariset mallit, kevät 2010

Harjoitus 4, viikko 16

1. Tarkastellaan lineaarista mallia $\boldsymbol{Y} \sim N_n(\boldsymbol{X\beta}, \sigma^2 I_n)$ $(\boldsymbol{\beta} \in \mathbb{R}^p, \sigma^2 > 0, \mathrm{r}(\boldsymbol{X}) = p)$. Osoita, että parametrivektorin $[\boldsymbol{\beta}' \ \sigma^2]'$ Fisherin informaatiomatriisi
$$\boldsymbol{i}(\boldsymbol{\beta}, \sigma^2) = E \begin{pmatrix} -\partial^2 l(\boldsymbol{\beta}, \sigma^2; \boldsymbol{Y})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}' & -\partial^2 l(\boldsymbol{\beta}, \sigma^2; \boldsymbol{Y})/\partial\boldsymbol{\beta}\partial\sigma^2 \\ -\partial^2 l(\boldsymbol{\beta}, \sigma^2; \boldsymbol{Y})/\partial\sigma^2\partial\boldsymbol{\beta}' & -\partial^2 l(\boldsymbol{\beta}, \sigma^2; \boldsymbol{Y})/\partial\sigma^2\partial\sigma^2 \end{pmatrix}$$
on
$$\boldsymbol{i}(\boldsymbol{\beta}, \sigma^2) = \begin{pmatrix} \sigma^{-2}\boldsymbol{X}'\boldsymbol{X} & \mathbf{0} \\ \mathbf{0} & n/2\sigma^4 \end{pmatrix}.$$
(Vihje: Kannattaa ehkä laskea ensin $\partial l(\boldsymbol{\beta}, \sigma^2; \boldsymbol{y})/\partial\boldsymbol{\beta}$ kuten $\partial S(\boldsymbol{\beta})/\partial\boldsymbol{\beta}$ monisteen s. 7 ja sen jälkeen $\partial^2 l(\boldsymbol{\beta}, \sigma^2; \boldsymbol{y})/\partial\boldsymbol{\beta}\partial\sigma^2$ ja edelleen vastaavan satunnaisvektorin odotusarvo. $\partial^2 l(\boldsymbol{\beta}, \sigma^2; \boldsymbol{y})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'$ saadaan lasketuksi kuten $\partial^2 S(\boldsymbol{\beta})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'$ (ks. monisteen s. 8) ja $\partial^2 l(\boldsymbol{\beta}, \sigma^2; \boldsymbol{y})/\partial\sigma^2\partial\sigma^2$ vaatii pelkästään derivointia reaalisen muuttujan σ^2 suhteen. Huomaa myös symmetrisyys.)

2. Olkoon oikea (täysiasteinen) malli $\boldsymbol{Y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$ $(\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\boldsymbol{I}_n), \boldsymbol{\beta}_1 \in \mathbb{R}^{p_1}, \boldsymbol{\beta}_2 \in \mathbb{R}^{p_2}, \sigma^2 > 0)$. Oletetaan, että $\boldsymbol{\beta}_1$ estimoidaan kuitenkin käyttäen mallia, josta $\boldsymbol{X}_2$ on jätetty pois (eli malliyhtälö on $\boldsymbol{Y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_*$). Laske näin saadun $\boldsymbol{\beta}_1$:n PNS-estimaattorin odotusarvo ja selvitä myös sen todennäköisyysjakauma. Milloin tämä estimaattori on harhaton?

3. (Jatkoa harjoituksen 2 tehtävälle 1) (i) Johda tehtävän varianssianalyysimallissa parametrien $\mu_1, \ldots, \mu_p$ PNS-estimaattien lausekkeet normaaliyhtälöiden ratkaisukaavaa käyttäen. (ii) Johda PNS-estimaattorien odotusarvot, varianssit ja kovarianssit (eli odotusarvovektori ja kovarianssimatriisi) ja osoita, että PNS-estimaattori $\hat{\boldsymbol{\mu}} = [\hat{\mu}_1 \cdots \hat{\mu}_p]'$ noudattaa multinormaalijakaumaa.

4. Tarkastellaan kahta riippumatonta lineaarista mallia
$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \ \boldsymbol{\varepsilon}_i \sim N_{n_i}(\mathbf{0}, \sigma^2\boldsymbol{I}_{n_i}), \ \boldsymbol{\varepsilon}_1 \perp\!\!\!\perp \boldsymbol{\varepsilon}_2, \ \boldsymbol{\beta}_i \in \mathbb{R}^p, \ \sigma^2 > 0, \ \mathrm{r}(\boldsymbol{X}_i) = p, i = 1, 2.$$
Muodosta näistä matriiseja käyttäen yksi malli ja esitä parametrin $\boldsymbol{\beta} = [\boldsymbol{\beta}_1' \ \boldsymbol{\beta}_2']'$ PNS-estimaattorin lauseke. Mikä on saadun PNS-estimaattorin jakauma? Aputulos: Olkoon $\boldsymbol{A}$ ja $\boldsymbol{B}$ epäsingulaarisia neliömatriiseja. Tällöin
$$\begin{pmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B}^{-1} \end{pmatrix}.$$

5. Tarkastellaan kahden riippumattoman normaalisen otoksen mallia
$$Y_1, \ldots, Y_n \ \perp\!\!\!\perp, \ Y_i \sim \begin{cases} N(\mu_1, \sigma^2), & \text{kun } i = 1, \ldots, n_1 \\ N(\mu_2, \sigma^2), & \text{kun } i = n_1 + 1, \ldots, n_1 + n_2 = n \end{cases}$$
$(\mu_1, \mu_2 \in \mathbb{R}, \sigma^2 > 0, n_1, n2 > 1)$. Estimoi parametrit μ_1 ja μ_2 ehdolla $\mu_1 = \mu_2$ käyttäen monisteessa (s. 16) esitettyä rajoitetun PNS-estimaattorin kaavaa (2.8).